Value of information – SITEX Data analysis

Shubha Kadambe

(310)317-5755

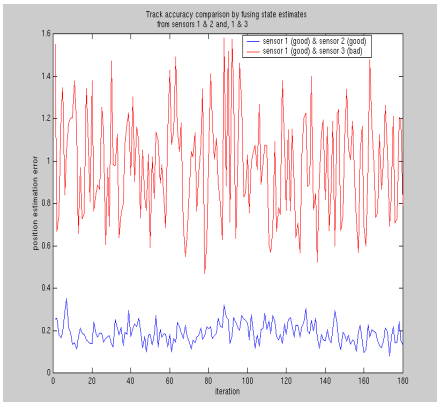
skadambe@hrl.com

Information Sciences Laboratory

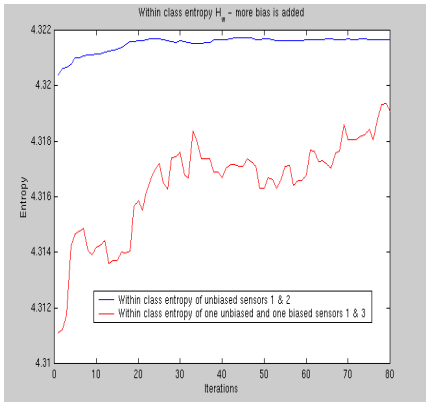
HRL Labs

3011 Malibu Canyon Rd., Malibu CA

Value of information – SITEX Data analysis


Track accuracy comparison by fusing state estimates from sensors 1 & 2 and, 1 & 3
sensor 1 (good) & sensor 2 (good)
sensor 1 (good) & sensor 3 (bad)
position estimation error
iteration

Mutual Information Metric used to determine best combination (blue) of sensors to fuse.


Within class entropy H_m - more bias is added
Entropy
Iterations
Within class entropy of unbiased sensors 1 & 2
Within class entropy of one unbiased and one biased sensors 1 & 3

Within Class entropy Metric used to discriminate biased sensor (red) vs unbiased sensor (blue)

New Ideas

- Theoretical performance analysis of
  - Detectors, trackers and classifiers in a network of distributed sensors
- Information theoretic based metrics for the performance analysis
- Lower and upper bound of performance using
  - Information from single sensor/node and decisions from the neighboring nodes
  - Information from multiple sensors/node and decisions from the neighboring nodes

Impact

- Theoretical framework for assessing the decision accuracy in a network of distributed sensors
- Determining optimal performance of algorithms under different conditions
- Enable development of optimal and robust algorithms

Schedule

- Development of Information theoretic metrics
- Development of lower bound
- Development of upper bound
- Performance analysis of algorithms
- Extraction of robust features
- Markov-model based robust classifier
- Kalman filter based robust tracker
- CDWR based robust detector

10/18/2001(start) | 9/17/2002 | 9/2003

Information theoretic based metrics: Conditional entropy & Mutual information

- Entropy is a measure of uncertainty.
- Let $H(x)$ be the entropy of previously observed events.
- Let y be the estimated features from another sensor which can be looked at as a new set of events.
- We can measure the uncertainty of x after observing y by using the conditional entropy which is defined as $H(x|y) = H(x,y) - H(y)$ with the property $0 \leq H(x|y) \leq H(x)$.
- Here, $H(x,y)$ is the joint entropy of observations x and y.
- The conditional entropy $H(x|y)$ represents the amount of uncertainty remaining about x after y has been observed.
- If the uncertainty is reduced then there is information gained by observing y.
- Therefore, we can measure the relevance of y by using conditional entropy.
- Another measure that is related to conditional entropy is mutual information $I(x,y)$ which is a measure of uncertainty that is resolved by observing y and is defined: $I(x,y) = H(x) - H(x|y)$

Mutual information as a measure of accuracy

- Let A = {a_k} k = 1, 2,... & B = {b_l} l = 1, 2,... be the set of features from sensor 1& 2
- Let p(a_i) be the probability of feature a_i.
- Let H(A), H(B) and H(A|B) be the entropy corresponding to sensor 1, sensor 2 and sensor 1 given sensor 2, respectively, and they are defined as:

$$H(A) = \sum_k p(a_k) \log \left[\frac{1}{p(a_k)} \right],$$

$$H(A|B) = H(A,B) - H(B) = \sum_l p(b_l) H(A|b_l) = \sum_l p(b_l) \sum_k p(a_k|b_l) \log \left[\frac{1}{p(a_k|b_l)} \right]$$

- The mutual information which is defined as I(A, B) = H(A) – H(A|B) corresponds to uncertainty that is resolved by observing B in other words features from sensor 2.
- Let us consider two types of sensors at node 2. Let the set of features of these two sensors be B_1 and B_2, respectively.
- If H(A|B_1) < H(A|B_2) then I(A, B_1) > I(A, B_2). This implies that the uncertainty is better resolved by observing B_1 as compared to B_2.
- This further implies that B_1 corresponds to relevant features and thus helps in improving the decision accuracy of sensor 1
- B_2 corresponds to non-relevant features with respect to sensor 1 and hence, B_2 should not be considered.

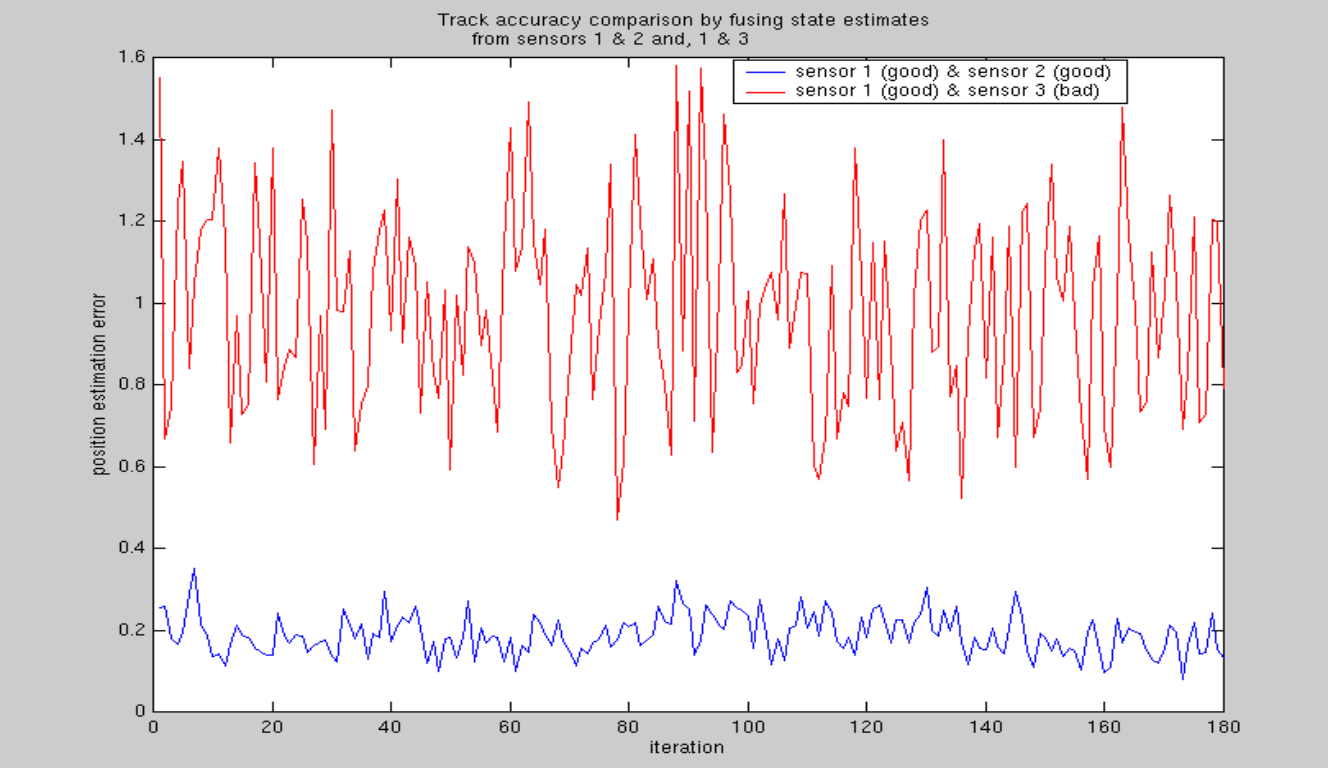
Mutual information metric in sensor fusion

- A network of radar sensors is used for tracking multiple targets.
- For tracking Kalman filter based approach is used.
- Each sensor node has a local and global Kalman filter based trackers.
- These target trackers estimate the target states - position and velocity in Cartesian co-ordinate system.
- The local tracker uses the local radar sensor measurements to estimate the state estimates while the global tracker fuses target states obtained from other sensors if it improves the accuracy of the target tracks.
- For this purpose mutual information metric was used.
- In the simulation
  - A network of three radar sensors and a single moving target with constant velocity were considered.
  - Two sensors were considered as good and one as bad.
  - Bad sensor - measurements were corrupted with high noise (e.g., SNR = -6 dB).
  - In this example the SNR of a good sensor is 10 dB.
  - The measurements from a radar at each sensor node was used to estimate the target states using the local Kalman filter algorithm.
  - The estimated target states at each sensor node were transmitted to other nodes

Mutual information metric in sensor fusion

- We consider the estimated state vector as the set of feature vector
- The mutual information metric based algorithm was implemented at sensor node 1 with the assumption it is a good sensor.
- Let the state estimate outputs of this node be A_g.
- Let the state estimate outputs of a second sensor correspond to B_g and a third sensor correspond to B_b.
- Entropy, conditional entropy and mutual information were computed
- If $I(A_g, B_g) > I(A_g, B_b)$ then the state estimates B_g was fused with A_g using the global Kalaman filter algorithm.
- Position estimation error was computed by comparing the fused state estimate with the true position.
- To compare the track accuracies, the state estimates from B_b and A_g were also fused using the global Kalman filter algorithm.
- The position estimation error was then computed the same way as explained above.

Mutual information metric in sensor fusion



- From this figure, it can be seen that the track accuracy after fusing state estimates from good sensors (1 & 2) is much better than fusing state estimates from a good sensor and a bad sensor (1 & 3). This implies that better mutual information correlates with better track accuracy.

Information theoretic metric: Within class entropy - measure of consistency

- Let there are N events (values) that can be classified in to m classes
- Let an event x_{ij} be the j^{th} member of i^{th} class where i = 1,2,.....,m, j = 1,2,.....,n_i and $\sum_{i=1}^{m} n_i = N$
- The entropy for this classification is:

$$H = \sum_{i=1}^{m}\sum_{j=1}^{n_i} p(i)\,p(x_{ij})\log\left[\frac{1}{p(i)\,p(x_{ij})}\right] = -\sum_{i=1}^{m}\sum_{j=1}^{n_i} p(i)\,p(x_{ij})\log\left[p(i)\,p(x_{ij})\right]$$

$$= -\sum_{i=1}^{m} p(i)\sum_{j=1}^{n_i} p(x_{ij})\log(p(x_{ij})) - \sum_{i=1}^{m} p(i)\log(p(i))\sum_{j=1}^{n_i} p(x_{ij})$$

$$= \sum_{i=1}^{m} p(i)H_i - \sum_{i=1}^{m} p(i)\log(p(i))$$

since $\sum_{j=1}^{n_i} p(x_{ij})\log(p(x_{ij}))$ is the entropy of a class i & $\sum_{j=1}^{n_i} p(x_{ij}) = 1$

$= H_w + H_b$ where H_w is called the entropy within classes and H_b is called the entropy between classes.

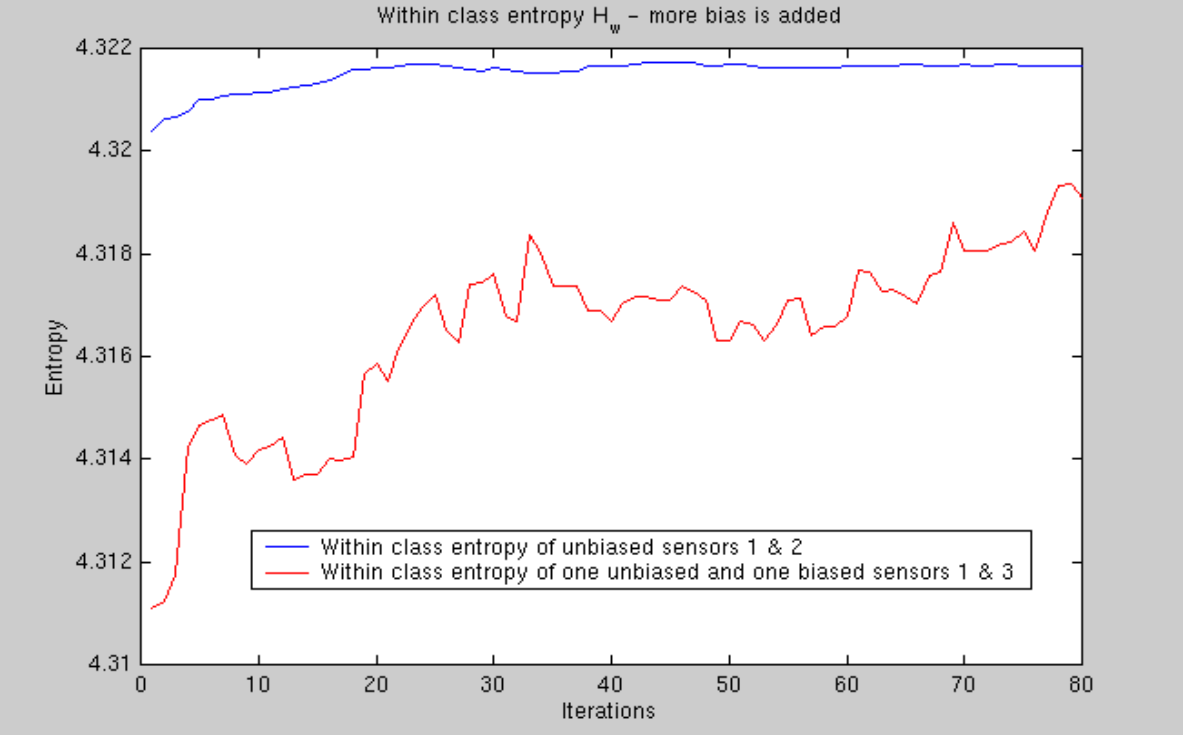
Within class entropy - measure of consistency

- The entropy H_w
  - is high if the values or events belonging to a class represent similar information and
  - is low if they represent dissimilar information.
  - This means H_w can be used as a measure to define consistency.
  - That is, if two or more sensor measurements are similar then its H_w is greater than if they are dissimilar.

Sensor discrimination using within class entropy metric

- The consistency measure was applied to discriminate between biased and unbiased sensors.
- In the simulations, the bias at one of the sensors was introduced as the addition of a random number to the true position of a target.
- The bias was introduced this way because the biases in azimuth and range associated with a radar sensor translate into measured target position that is different from the true target position.
- In addition, currently in our simulations, we are assuming that the sensors are measuring the target's position in the Cartesian co-ordinate system instead of the polar co-ordinate system.
- We considered three sensors – two were not biased and one was biased.
- The amount of bias was varied by multiplying the random number by a constant k i.e., measured position = (true position + k * randn) + measurement noise.

Sensor discrimination example



Plot of within class entropy of sensors 1 & 2 (unbiased sensors) and, 1 (unbiased) and 3 (biased). Bias constant k = 2

From this figure

- it can be seen that the within class entropy is greater when the two sensors are unbiased as compared to the within class entropy when one of them is biased.
- This indicates that the within class entropy can be used as a consistency measure to discriminate between sensors.

SITEX data analysis using Information theoretic metrics

- Based on the promising preliminary results, we believe that
  - information theoretic based metrics can be used in the theoretical performance analysis of detection, tracking and classification algorithms.
- Therefore, we further develop the information theoretic based metrics and apply them for SITEX data analysis.
- First, we use
  - the mutual information metric to test whether the new information help in improving the decision accuracy of the current node.
  - the consistency metric to decide whether the new information is consistent with the current node.
    - ✦This also helps in determining whether the sensor is functional or how much to weigh the decision of a neighboring node – useful for fusion or automatic clustering of sensors.

SITEX data analysis - bounds

- Lower bound
  - One sensor's information from each node but fuse only the decisions from the neighboring nodes
- Upper bound
  - fusion of information from all sensors on a node and also fusion of decisions obtained from other nodes

SITEX data analysis - status

- Identified the classifier and the detector for the initial analysis
- Currently both BAE's wideband and SITEX00 data is being used
- Data is being analyzed by
  - extracting features
  - and first computing the mutual information and within class entropy metrics
- For fusion Bayesian approach is being used
- Lower bound is being computed